Application of Several Visual Analytic Approaches in Detecting Technology Front

Luan Chunjuan

School of Public Administration & Law,WISE Lab, Dalian University of Technology
Linggong Road, Dalian, P. R. China

julielcj@163.com

***Abstract*-It is of great significance for technology analysists, technology policy makers and R&D managers to grasp technology front by applying information visualization technique via patent metric. Visualization analysis is widely applied in exploring scientific front, but seldom used in detecting technology front. In this article, the author tries three visual analytic approaches to detecting technology front by using patent data retrieved from the authoritative patent database worldwide of Derwent Innovations Index (abbr. DII) in the field of global aerospace field in 2009. The results show that different visual analytic approaches used in the paper can be applied to detect technology front; the technology fields of Communications and Computing & Control, especially technology of Communications is the highlight of global aerospace field in 2009; the clusters are obvious in the three visual analytic maps drawn by applying three different visual analytic approaches; each of the visual analytic approach has its own advantages and disadvantages. It is still a macro-study of this paper, more concrete analysis on detecting technology front is expected to explore in the near future.**

Keywords-technology front; visual analytic approaches; social network analysis; Ucinet; CiteSpace; VOSviewer

I. INTRODUCTION

Price stated that the first classic scientists should pay more attention to the latest work done by their colleagues, for research front often depends on the latest research findings [1]. Here "research front" should be understood as including both "scientific research front" which can be abbreviated as "science front", and "technological research front" which can be abbreviated as "technology front" or "technology frontier".

Technology front can be defined as the highlight of the latest research findings of technology research and development (abbr. R&D) [2]-[3]. Some studies on technology front were explored by scholars [2], [4-6]. But little studies on technology front were done by using visual analytic approaches.

Visualization analysis is popularly applied in detecting scientific research front. With the development of the research of information visualization, many kinds of visual software have been developed and applied in STS research area and scientometric research field. Visualization analysis in scientific frontier studies mainly includes citation analysis and co-occurrence analysis. Small & Sweeney et al., Leydesdorff, Tijssen and Vanleeuwen, Chen & Paul et al., McCain & Verner et al., Leydesdorff, Zou & Wu et al., have applied co-citation analysis mapping science [7]-[13]; Van Den Besselaar & Heimeriks, Zhang & Wolfram et al, have used co-occurrence analysis mapping science [14]-[15]. But visualization analysis has been little used in detecting technology front so far, and the existing studies on detecting technology front has been mainly explored by using social network analysis (abbr. SNA) related to keyword-analysis or co-citation analysis [16]-[17].

In this study, we should introduce several visual analytic approaches the author familiar with in detecting technology front using patent data in the technology field of Aircraft, Aviation & Cosmonautics globally in 2009, and the visual analytic approaches include SNA analysis, Cite Space analysis, VOS viewer analysis.

II. DATA COLLECTION AND RESEARCH METHODOLOGY

Patent as one of the most important technology resources covers more than 90% of R&D outputs [18]-[19], and patent is widely applied as an important index to explore technology front in analyzing the latest patent literature. So it is concluded that the latest patent data can be used to analyse technology front. The motivation of this research is to help researchers grasp the trend and the art of a certain technology area to make correct R&D decisions.

The study is based on bibliographic data retrieved from DII (Derwent Innovations Index), one of the most authoritative databases collecting patent documents in the world, begun in 1963 and currently published by the Thomson Reuters firm. DII includes three parts: Chemistry, Engineering and Electric & Electronics. Every week 25, 000 patent documents published by more than 40 patent offices and 45, 000 patent citation documents from 6 important copyright offices are input into DII. The date used in our statistics is of publication.

The data search path operates with "B64*" in the International Patent Classification Code (abbr. IPC), B64 represents the technology field of "Aircraft, Aviation & Cosmonautics" (abbr. "aerospace field"), i.e. IPC=B64= aerospace field; time span=2009; and database=Derwent Innovations Index. The data was retrieved on March 20, 2010, and we downloaded all the 3660 patent data in global aerospace field in 2009.

Technology field occurrence analysis is applied in the study. Derwent Manual Code (abbr. DMC) is chosen as technology field classification, for DMC is a kind of classification with details indicated by Derwent technicians, and DMC is considered more accurate in classification. Several visual analytic approaches would be applied in the study, including SNA (Ucinet) [20], CiteSpace[21]-[22] and VOSviewer (abbr. VOS) [23].

III. APPLICATION OF SNA (UCINET) IN DETECTING TECHNOLOGY FRONT

SNA is an important methodology in the sociological studies in the west. It was firstly put forward by Brown, a famous humanist in Britain. With the development of the

theory, methodology and technique, SNA has gradually become a significant research paradigm in social structure studies, and some network analysis softwares emerge as the requirement of times, such as Pajek [24], Ucinet [20], NWB [25] et al.. SNA is widely used in interpersonal relationship studies [26]-[29]. But little studies on technology front have been done by applying SNA. Shih and Chang [30] investigated the structure of international technology diffusion by using network analysis. This study identifies two types of catch-up strategies that newly industrialized or developing countries can use to move up the international technology stratification. Mika & Elfring et al. [31] provides a methodological innovation based on semantic technology for dealing with heterogeneity in electronic data sources. By using methods of network analysis, they confirm the effect of Structural Holes and provide novel explanations of scientific performance based on cognitive diversity in social networks.

DMC was adopted to analyze the technology front of global aerospace field in 2009. Using the software of Bibexcel developed by Persson [32], we got 1435 different technology fields, and then we chose top 87 fields by their frequency, by adopting the drawing tool Netdraw of Ucinet, we drew Figure 1.

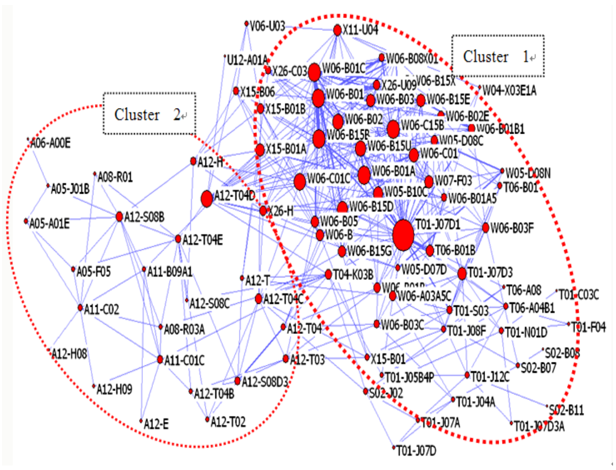

Fig. 1 DMC co-occurrence network of global aerospace field frontier in 2009 by Ucinet (degree centrality)

W: Communications; T: Computing and Control; A：Plasdoc

It is clearly that there are two comparatively independent parts in Figure 1, and they are outlined separately with dash lines. The right part (cluster 1) is a more intensive part, and it is mainly composed of technology W and T, especially W, i.e. technology of Communications; the left part (cluster 2) appears to be more loose, and it is mainly composed of technology A, i.e. Plasdoc technology. Figure 1 shows that technology of Communications and Computing & Control is the highlight of global aerospace field frontier in 2009. It can be concluded from this test that SNA (Ucinet) can be used to detect technology front by mapping technology network in which technology fields with close relationship are tend to form an independent technology cluster.

IV. APPLICATION OF CITESPACE IN DETECTING TECHNOLOGY FRONT

Cite Space is a freely available Java application for visualizing and analyzing trends and patterns in scientific literature developed by Chaomei Chen. It is designed as a tool for progressive knowledge domain visualization [33]. It focuses on finding critical points in the development of a field or a domain, especially intellectual turning points and pivotal points. Detailed case studies are given in the paper of "CiteSpace II: Detecting and visualizing emerging trends and transient patterns in scientific literature" [21] and other publications. CiteSpace provides various functions to facilitate the understanding and interpretation of network patterns and historical patterns, including identifying the fast-growth topical areas, finding citation hotspots in the land of publications, decomposing a network into clusters, automatic labelling clusters with terms from citing articles, geospatial patterns of collaboration, and unique areas of international collaboration. CiteSpace supports structural and temporal analyses of a variety of networks derived from scientific publications, including collaboration networks, author co-citation networks, and document co-citation networks. It also supports networks of hybrid node types such as terms, institutions, and countries, and hybrid link types such as co-citation, co-occurrence, and directed citing links [34].

Here the function of co-occurrence of CiteSpace was applied to detect the technology front of global aerospace field in 2009. Firstly, we should use the data-transferred-function of CiteSpace to transfer the patent data downloaded from DII into the web of science format, for the primary source of input data for CiteSpace is the Web of Science, though it also provides some simple interfaces for obtaining data from PubMed, arXiv, ADS, and NSF Award Abstracts [35].

We still choose DMC as the technology classification in the following study, that is, the node in the map represents DMC, a kind of patent classification given by the technicians of Thomson Reuters Corp...After adjust the threshold value as (2, 1, 10), (2, 1, 10), (2, 1, 10), and set the slice=1, then go Cite Space, and then we get the figure 2. 114 nodes which are included in Figure 2.

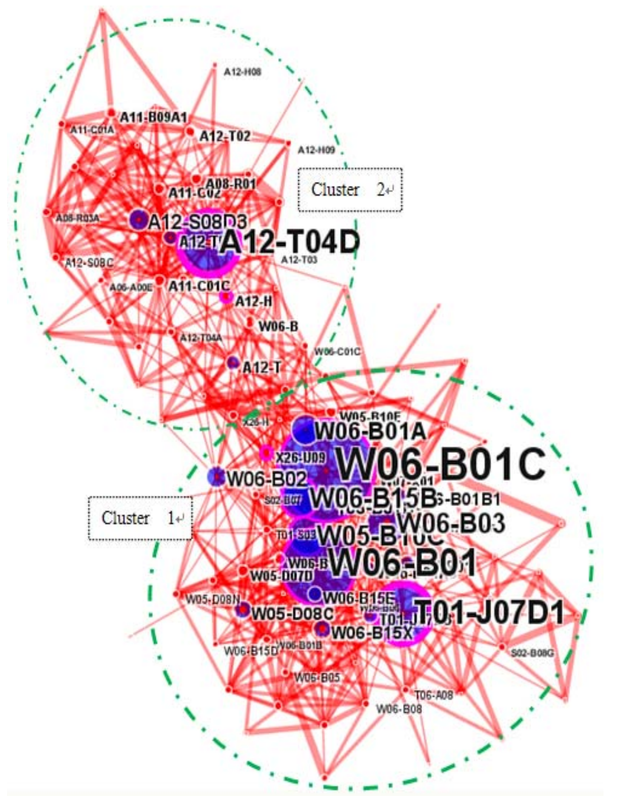

Fig. 2 DMC co-occurrence network of global aerospace field frontier in 2009 by CiteSpace (node frequency)
W: Communications; T: Computing and Control; A：Plasdoc

It is apparently that Figure 2 includes 2 comparatively independent components which at the same time are linked with each other: Cluster 1 and Cluster 2 which outlined with dash lines by the author. Cluster 1 mainly consists of technology /field of W and technology field of T, that is, Cluster 1 includes the technology of Communications and the technology of Computing & Control. Cluster 2 mainly consists of technology field of A, i.e. technology of Plasdoc. Figure 2 also shows that, cluster 1 is the main part of the map, indicating that the technology of Communications and the technology of Computing & Control is the highlight of global aerospace field frontier in 2009. It can help us identify technology front clusters by using software of CiteSpace. Technology fields with near relationship always form into a comparative separate technology cluster, and different technology clusters show us the main areas of R&D focuses.

V. APPLICATION OF VOS IN DETECTING TECHNOLOGY FRONT

VOSviewer (abbr. VOS) has been developed by Nees Jan van Eck and Ludo Waltman, who are two researchers working at the Centre for Science and Technology Studies, Leiden University, since June, 2009. VOS is a freely available computer program that can be used to construct maps of authors or journals based on co-citation data or to construct maps of keywords based on co-occurrence data. [23]. VOS, which stands for visualization of similarities, is a method that visualizes similarities among objects. The aim of VOS is to provide a low-dimensional visualization in which objects are located in such a way that the distance between any pair of objects reflects their similarity as accurately as possible. Objects that have a high similarity should be located close to each other, whereas objects that have a low similarity should be located far from each other. VOS can be seen as an alternative to multidimensional scaling, which is the standard approach to visualize similarities between objects [36].

A net document with 419 nodes was chosen to draw VOS map (See Figure 3).

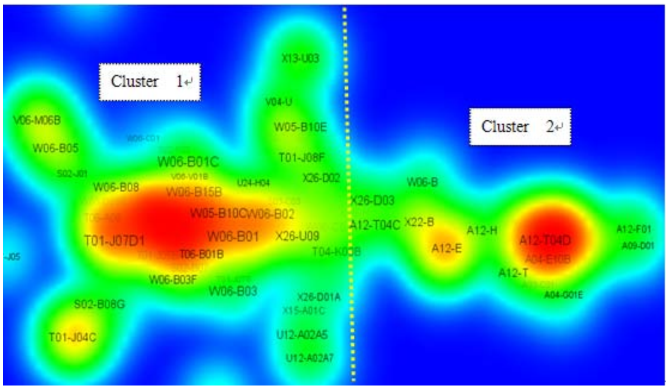

Fig. 3 DMC co-occurrence network of global aerospace field frontier in 2009 by VOSviewer (frequency)

W: Communications; T: Computing and Control; A：Plasdoc

There are two separately centralities which are divided into two parts with the dotted line by the author in Figure 3: a main centrality in the left of the map and a minor centrality in the right of the map. The left part (Cluster 1) mainly consists of the technology field of W and T, i. e. technology of Communications and technology of Computing & Control; while the right part (Cluster 1) mainly consists of technology field of A, i. e. technology of Plasdoc. Figure 3 also shows that Cluster 1 is the main body of the map, that is, technology of Communications and Computing & Control is the highlight of global aerospace field frontier in 2009. This test by using VOS software tells us that those technology fields with close relationship are always forming into a same cluster intensively, so we can conclude that VOS software can be used to detect technology front.

VI. CONCLUSIONS

•From the above analyses, we can conclude that different visual analytic approaches used in the paper can be applied to detect technology front.

• According to the case study, it is found that the technology fields of Communications and Computing & Control, especially technology of Communications is the highlight of global aerospace field in 2009.

•The clusters are obvious in the three maps, though map 1 is drew by “degree centrality”, whereas map 2 & map 3 are drew by “frequency”. Whether “degree centrality” or “frequency” all react amount of relationship between one node and other nodes in the maps, i. e., they almost have same meaning in the maps, and just due to different software has different indicators.

•In each map, Cluster 1 is the main body in the map, and it represents the technology of Communications and Computing & Control; Cluster 2 is the minor body in the map, and it represents the technology of Plasdoc.

•Comparing the three visual analytic approaches, each of them has its own advantages and disadvantages. Figure 1, drawn by Ucinet, is quite clear, but it can only include no more than 100 nodes to keep the map and the label clear; clusters in figure 2 are very ideal, and this map drawn by using CiteSpace can includes more than 100 nodes, but it is different from making each node and its label clearly; when it comes to figure 3 applied VOS, the biggest advantage of it is that it can include much more nodes, hundreds, even thousands according to Nees Jan van Eck & Ludo Waltman [36]. Another advantage of figure 3 is that the map well expresses the highlight in red color by density. The disadvantage of figure 3 is that the link relationships and their strengths of the nodes can’t be identified for lack of link lines.

ACKNOWLEDGMENT

The research was supported by the Natural Science Foundation of China (NSFC) under Grant 70773015; Project of the Ministry of Education Humanities & Social Sciences Research Foundation under Grant 09YJA630014; Project of Graduate School, DUT under Grant JG1048; Project of Guangdong Natural Science Foundation under Grant 10451805702006271; as well as Project of Liaoning Province Social Sciences Foundation under Grant L08BJY062.

REFERENCES

[1] D.D price. Networks of scientific papers. Science，1965，(149)：510-515.

[2] M. A. Qureshi，V. P. Kharbanda. CHINAS ENERGY-POLICY - PROGRESS ON THE BIOGAS TECHNOLOGY FRONT. Journal of Scientific & Industrial Research，1982，41(7)：413-416.

[3] L. Iacovone，G. A. Crespi. Catching up with the technological frontier:Micro-level evidence on growth and convergence. Industrial and Corporate Change，1960，19(6)：2073-2096.

[4] J. Ha，Y. J. Kim，J. W. Lee. Optimal Structure of Technology Adoption and Creation: Basic versus Development Research in Relation to the Distance from the Technological Frontier. Asian Economic Journal，2009，23(3)：373-395.

[5] F. Caselli，W. J. Coleman. The US technology frontier. American Economic Review，2002，92(2)：148-152.

[6] F. Caselli，W. J. Coleman. The world technology frontier. American Economic Review，2006，96(3)：499-522.

[7] L. Leydesdorff. THE GENERATION OF AGGREGATED JOURNAL-JOURNAL CITATION MAPS ON THE BASIS OF THE CD-ROM VERSION OF THE SCIENCE-CITATION-INDEX. Scientometrics，1994，31(1)：59-84.

[8] H. Small，E. Sweeney，E. Greenlee. CLUSTERING THE SCIENCE CITATION INDEX USING CO-CITATIONS .2. MAPPING SCIENCE. Scientometrics，1985，8(5-6)：321-340.

[9] R. J. W. Tijssen，T. N. Vanleeuwen. ON GENERALIZING SCIENTOMETRIC JOURNAL MAPPING BEYOND ISIS JOURNAL AND CITATION DATABASES. Scientometrics，1995，33(1)：93-116.

[10] C. M. Chen，R. J. Paul，B. O'Keefe. Fitting the jigsaw of citation: Information visualization in domain analysis. Journal of the American Society for Information Science and Technology，2001，52(4)：315-330.

[11] K. W. McCain，J. M. Verner，G. W. Hislop，et al. The use of bibliometric and Knowledge Elicitation techniques to map a knowledge domain: Software Engineering in the 1990s. Scientometrics，2005，65(1)：131-144.

[12] L. Leydesdorff. Mapping interdisciplinarity at the interfaces between the science citation index and the social science citation index. Scientometrics，2007，71(3)：391-405.

[13] F,Zou,M. X. Wu，K. L. Wu. Outcomes associated with ophthalmology, optometry and visual science literature in the Science Citation Index from mainland China, 2000-2007. Scientometrics，2009，81(3)：671-682.

[14] P. Van Den Besselaar，G. Heimeriks. Mapping research topics using word-reference co-occurrences: A method and an exploratory case study. Scientometrics，2006，68(3)：377-393.

[15] J. Zhang，D. Wolfram，P. L. Wang，et al. Visualization of health-subject analysis based on query term co-occurrences. Journal of the American Society for Information Science and Technology，2008，59(12)：1933-1947.

[16] B. Lee，Y. I. Jeong. Mapping Korea's national R&D domain of robot technology by using the co-word analysis. Scientometrics，2008，77(1)：3-19.

[17] M. Igami. Exploration of the evolution of nanotechnology via mapping of patent applications. Scientometrics，2008，77(2)：289-308.

[18] Yang J Liu C. Decoding Patent Information Using Patent Maps. Data Science Journal，2008，7(0)：14-22.

[19] X. W. Wang. Patent Mapping of Technology Competition among Fortune 500 Enterprises Based on Global Patent Co-citation Analysis. Research and development management，2010，31(2)：19-25.

[20] B. W. Cheng, Y. L. Wang etc. Logarithmic normal distribution of S&T scale merit. Information studies: theory and application，2000，23(1)：18-23.

[21] C. M. Chen. CiteSpace II: Detecting and visualizing emerging trends and transient patterns in scientific literature. Journal of the American Society for Information Science and Technology，2006，57(3)：359-377.

[22] Chen C. Searching for intellectual turning points: Progressive knowledge domain visualization. Proceedings of the National Academy of Sciences，2004，101 (suppl-1)：5303-5310.

[23] Nees Jan van Eck，Ludo Waltman. VOSviewer. 2010. http:/ /www. vosviewer.com/.

[24] R. Stone，J. A. C. Brown. OUTPUT AND INVESTMENT FOR EXPONENTIAL-GROWTH IN CONSUMPTION. Review of Economic Studies，1962，29(80)：241-245.

[25] Y. P. Li. On evaluation of Price exponential growth model. Information studies: theory and application，1999，22(1)：11-13.

[26] Milgram S. The Small World Problem. Psychology Today，1967，(2).

[27] Y. Park，W. M. Zhu. Social network analysis for studying complex social network phenomena. Research Quarterly for Exercise and Sport，2007，78(1)：A8-A9.

[28] N. R. Hassan. Using Social Network Analysis to Measure IT-Enabled Business Process Performance. Information Systems Management，2009，26(1)：61-76.

[29] P. Hawe，L. Ghali. Use of social network analysis to map the social relationships of staff and teachers at school. Health Education Research，2008，23(1)：62-69.

[30] H. Y. Shih，T. L. S. Chang. International diffusion of embodied and disembodied technology: A network analysis approach. Technological Forecasting and Social Change，2009，76(6)：821-834.

[31] P. Mika，T. Elfring，P. Groenewegen. Application of semantic technology for social network analysis in the sciences. Scientometrics，2006，68(1)：3-27.

[32] Olle Persson. A tool-box developed by Olle Persson, Inforsk, Umeå univ, Sweden http://www8.umu.se/inforsk/Bibexcel/.

[33] C. M. Chen，J. Kuljis. The rising landscape: A visual exploration of superstring revolutions in physics. Journal of the American Society for Information Science and Technology，2003，54(5)：435-446.

[34] C.M.Chen.CiteSpace.2010.http://cluster.cis.drexel.edu/~cchen/citespace/.

[35] Chaomei Chen. CiteSpace: http://cluster.cis.drexel.edu/~cchen /citespace/.2011-08-01.

[36] Ludo Waltman Nees Jan van Eck. VOS. http://www.neesjanvaneck.nl /vos/. 2088-08-02.